VOLUME 76 SEPARATE No. D-XVI

PROCEEDINGS

AMERICAN SOCIETY OF CIVIL ENGINEERS

JULY, 1950

AMERICAN SOCIETY OF CIVIL ENGINEERS FOUNDED 1852

DISCUSSION OF STREAM FLOW VARIABILITY

(Published in September, 1949)

By Ven Te Chow, Dana M. Wood, H. Alden Foster, C. S. Ospina and G. Tama, Raphael G. Kazmann, Jack Bruin and H. E. Hudson, Jr., H. W. Lull, William D. Mitchell, C. R. Ownbey, S. P. Wing, J. E. McKee, G. B. Schroyer and E. T. Schuleen, and E. W. Lane

HYDRAULICS DIVISION

Headquarters of the Society
33 W. 39th St.
New York 18, N.Y.

PRICE $0.50 PER COPY

Published at Prince and Lemon Streets, Lancaster, Pa., by the American Society of Civil Engineers. Editorial and General Offices at 33 West Thirty-ninth Street, New York 18, N. Y. Reprints from this publication may be made on condition that the full title of paper, name of author, page reference, and date of publication by the Society are given.

DISCUSSION

VEN TE CHOW.[21]—Following the computation procedure proposed in the paper, the variability index can be expressed by the formula:

$$I_v = \sqrt{\frac{\Sigma\,(y - \bar{y})^2}{n - 1}} \quad \text{(1)}$$

in which y is the logarithm of the selected discharge at a 10% interval of the duration curve; $\bar{y}$ is the mean of y; $y - \bar{y}$ is the deviation from the mean; and n is the number of selected discharges, or $n = 10$ in the given procedure. In a practical operation, it is seldom convenient to take the deviations from the actual mean, since such deviations usually involve decimals which are cumbersome to handle when squared. For greatest convenience (especially in machine computation), another form of Eq. 1 is obtained by utilizing the mathematical principle that the sum of the squares of deviation from the mean is equal to the sum of the squares of deviation from zero, less the product of the total and the mean. Modified slightly, this principle is expressed as

$$I_v = \sqrt{\frac{n}{n - 1}\,(\overline{y^2} - \bar{y}^2)} \quad \text{(2)}$$

Eq. 2 is known as the Gaussian formula, which is used to estimate the standard deviation of population in statistics. In Eq. 2 $\overline{y^2}$ is the mean of the squares of y. The variability index can be defined mathematically and simply by this formula.

The numerical method of computing the variability index was primarily presented for the purpose of eliminating the effects introduced by variations in personal judgment. However, as the method used the values of discharges read from the duration curve as plotted, it did not eliminate the personal factor introduced in plotting the curve. A complete elimination of the influence of personal judgment is possible only by replacing the "eye fitting" by a mathematical process—that is, the values of $\bar{y}$ and $\overline{y^2}$ contained in Eq. 2 should be computed directly from the raw data. Such data can be grouped by the method of statistical classification to reduce the number of data handled in computation. If necessary, the computed standard deviation can be corrected by applying the Sheppard correction.

Since the mathematical part of the paper has been omitted, the method of producing a duration curve from the given values of the variability index and

NOTE.—This paper by E. W. Lane and Kai Lei was published in September, 1949, *Proceedings*. The numbering of footnotes, tables, equations, and illustrations in this Separate is a continuation of the consecutive numbering used in the original paper.

[21] Hydr. Eng. Laboratory, Univ. of Illinois, Urbana, Ill.

the mean annual runoff cannot be checked readily. However, the stipulation (under the heading, "Use of the Variability Index") that—

> "* * * the ratio of the discharge exceeded 15.87% of the time to the discharge exceeded 50% of the time was equal to the variability index selected"

—would make the variability index greater than unity for most streams, because the discharge that is exceeded 15.87% of the time is usually greater than the discharge exceeded 50% of the time. Apparently this estimate does not agree with the values of the variability index listed in Table 1.

The knowledge of hydrologic statistics has been developed and extended greatly in recent years. Many investigators[17,22] have found that the distribution of most hydrologic phenomena is skew. Probability plottings of the Hazen type were also found unsatisfactory for such frequencies. The computation of asymmetric probabilities is now possible in the field of hydrology; and the development of a special probability paper[16,23] opened the way for a straight-line plotting of the duration curve. Therefore the study of stream flow would be advanced if a suitable asymmetric probability plotting could be adapted to it. The values of discharge should be plotted on a linear scale, so that the variability index would be equal to the standard deviation, in a statistical sense.

Many hydrologists like to use the coefficient of variation, C_v, instead of the standard deviation, σ, for the study of variability. The coefficient of variation is equal to the standard deviation divided by the mean, or

$$C_v = \frac{\sigma}{\bar{y}} \qquad (3)$$

For purposes of comparison, the coefficient of variation is preferable to the standard deviation, because the former, being dimensionless, eliminates the effects of the difference from the mean. This elimination is desirable, especially for stream discharge in more arid regions where the annual rainfall tends to be smaller as well as more variable.

For any statistical study, the quantity of the available data is as important as the quality. As long as the quality is kept homogeneous, the longer record, or larger number of data, would make possible more accurate estimates of the frequencies of stream flow. W. D. Potter,[24] of the Soil Conservation Service, United States Department of Agriculture, has made a probability study of rainfall and runoff. The study was confined to annual maximum values; but it is worthy to note that the frequency of distribution is a function of years of record, which is a measure of the number of data used. Since the variability index is a measure of the slope of the duration curve (a frequency curve), it should be some function of the length of record. Therefore, the variability index would be of more statistical value if the two were compared on the basis of the same length of record.

[17] "An Asymmetric Probability Function," by J. J. Slade, *Transactions*, ASCE, Vol. 101, 1936, p. 35.

[22] "Floods Estimated by Probability Method," by E. J. Gumbel, *Engineering News-Record*, June 14, 1945, p. 97.

[16] "Straight Line Plotting of Skew Frequency Data," by R. D. Goodrich, *Transactions*, ASCE, Vol. 91, 1927, p. 1.

[23] "A Simple Method of Estimating Flood Frequency," by Ralph W. Powell, *Civil Engineering*, February, 1943, p. 105.

[24] "Simplification of the Gumbel Method for Computing Probability Curves," by W. D. Potter *No. SCS-TP-78*, Soil Conservation Service, U.S.D.A., Washington, D. C., May, 1949.

DANA M. WOOD,[25] M. ASCE.—The selection of an analytical method depends upon the use to which the results are to be put and the time available for the study. Cost is also a factor in any engineering office. Two suggestions concerning the duration curve may be of value to those desiring to extend this work.

(1) Instead of using a scale of "discharge in cubic feet per second per square mile," substituting "ratio to the average flow" for the period of record under investigation will result in relative variability curves of flow. The values of departures at 10% intervals can then be obtained similarly from this curve. This procedure eliminates the size of the drainage area and the magnitude of the runoff in any comparisons. A stream having a flow of 0.5 cu ft per sec per sq mile can be compared with one having a flow 2.5 cu ft per sec per sq mile with respect to the slopes of the duration curves—that is, the magnitude of the departures from the average, regardless of the size and other characteristics of the drainage area. The writer would like to know if this approach might in any way simplify the logarithmic calculations.

(2) The second point of importance is that the same period of years should be used when comparing different records in the same general locality. It is well known that the magnitude of the runoff varies in cycles of wet and dry years, although little enough is known about the amplitude and magnitude of the curve of variation. However, in the Tennessee basin and numerous other basins, the years from 1930 to 1948 have been generally in the "dry" part of a cycle. The previous nineteen years, from 1910 to 1929, were generally in a flat part of the cycle having normal annual variations. It is quite obvious that any analysis could lead to erroneous conclusions if records in the two separate periods were compared. This point is well illustrated by the doubt regarding the analysis of records in different periods of years when attempting to determine the frequency of occurrence of floods of great magnitude.[26] However, for more distant localities, where the runoff cycles coincidentally may be radically different, further clarification and the development of methods of adjustment are needed.

This paper offers suggestions for many interesting exploratory studies, especially for the younger engineers training in these fields who are not afraid of undertaking many laborious calculations.

H. ALDEN FOSTER,[27] M. ASCE.—The authors have presented a new and interesting method for classifying the daily-flow duration curves of various streams by a single "variability index," together with the mean flow of the stream. At first glance, the variability index might seem to be related in some way to the classification of skew probability curves by the coefficient of variation and the coefficient of skew, as originally proposed by the late Allen Hazen,[12] M. ASCE, and further developed by the writer.[15] However, the latter classification was intended to be used in the study of selected samples of the original

---

[25] Chf., Power Studies Branch, Div. of Water Control Planning, TVA, Knoxville, Tenn.

[26] "Possible and Probable Future Floods," by William P. Creager, *Civil Engineering*, November, 1939, p. 668.

[27] Prin. Associate, Parsons, Brinckerhoff, Hall & Macdonald, New York, N. Y.

[12] "Storage to Be Provided in Impounding Reservoirs for Municipal Water Supply," by Allen Hazen, *Transactions*, ASCE, Vol. LXXVII, December, 1914, p. 1539.

[15] "Theoretical Frequency Curves and Their Application to Engineering Problems," by H. Alden Foster, *ibid.*, Vol. LXXXVII, 1924, p. 142.

data, such as annual floods or mean annual runoff values, for which the number of items actually used was limited, so that the calculations could be performed without an excessive amount of work. To compute the coefficients of variation and skew for a daily-flow record of several years would be entirely impracticable.

Certain points related to the proposed method may need further clarification:

(1) The calculation of the variability index is independent of the unit of runoff used, whether cubic feet per second or cubic feet per second per square mile, or flow in terms of mean flow—because any change in the unit does not change the shape of the duration curve but only the relative values; and, when the curve is plotted on logarithmic probability paper, the result would be only to shift the entire curve parallel to itself.

(2) The variability index is a function only of the slope of the straight-line duration curve when plotted on logarithmic probability paper. Hence, the value of the index is not changed if the curve is shifted parallel to itself on the plotting paper.

(3) When applying the method for a study in which the data for plotting a daily-flow duration curve are incomplete, the following procedure should be followed—

(*a*) Assume a value for the variability index—I;

(*b*) Determine a value, Q_i, such that $\log Q_i = I$ (Q_i being a daily flow value, expressed in terms of the median stream flow);

(*c*) Assume a value for the median stream flow Q_{50} (the flow exceeded 50% of the time), which may be estimated from the mean flow, as explained subsequently;

(*d*) Compute the discharge exceeded 15.87% of the time as the product of Q_i and Q_{50} (derivation of this percentage value being given subsequently); and

(*e*) Then plot the duration curve as a straight line on the logarithmic probability paper, passing through the two stream flow values for 15.87% and 50% of time.

Probability Paper.—To clarify the foregoing procedure, some explanation of the method of constructing the logarithmic probability paper is necessary. As far as the writer knows, no adequate description of such a method of construction has been published. Construction of this plotting paper is based on the use of the "normal curve of error," which has the form:

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \qquad (4)$$

and is plotted in Fig. 4(*a*). The maximum ordinate of the curve is $y_o = h/\sqrt{\pi}$. The "probable error" is $r = 0.4769/h$. The probable error is also defined by the fact that the area under the curve of error between the abscissas $-r$ and $+r$ is one half of the total area under the curve. The normal curve of error is known as a "frequency curve." If the frequency curve is integrated, the result is a "duration curve," or in the present case the curve of the "probability integral," as shown in Fig. 4(*b*). When plotted in this way, the axis

of the ordinates becomes the "percentage-of-time" scale, commonly used in plotting duration curves. As shown in Fig. 4(*b*) the part of the duration curve between $x = -r$ and $x = +r$ is represented by the distance between 25% and 75% in the percentage-of-time scale.

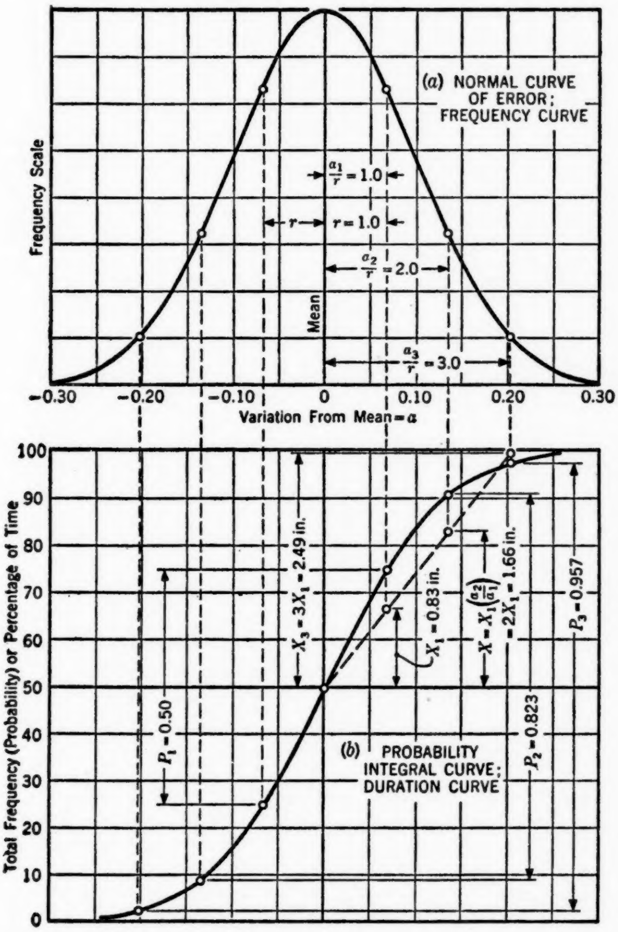

Fig. 4.—Relation of Probability Integral Curve to Probability Plotting Paper

If the abscissas, a, of the frequency curve are expressed in terms of r, the area under the curve from $-a/r$ to $+a/r$ may be found from values of the probability integral, as tabulated in mathematical textbooks. This integral may be expressed as follows:

$$P_a = \frac{h}{\sqrt{\pi}} \int_{-a}^{+a} e^{-h^2x^2}\, dx = \frac{2}{\sqrt{\pi}} \int_{o}^{ah} e^{-h^2x^2}\, d(h\,x) \qquad (5a)$$

or

$$P_a = \frac{2}{\sqrt{\pi}} \int_0^X e^{-x^2}\, dx \text{..................(5b)}$$

in which

$$X = a\,h = 0.4769\,\frac{a}{r} \text{......................(6)}$$

Tables of the probability integral generally give values of P_a as a function[28] of X or of[29] a/r. If P_a is the value of the probability integral for a "residual" of a/r (area from $-a/r$ to $+a/r$), then P_a will be the percentage-of-time interval on the duration curve between values of $(Q_{50} - a)$ and $(Q_{50} + a)$. The symbol Q_{50} denotes the value of the median, which is also the mean when the frequency curve is symmetrical.

If two stream flow values, Q_1 and Q_2, differ from the mean by a_1 and a_2, their percentage-of-time range on the duration curve (shown as P_1 and P_2 in Fig. 4(*b*)), or their values of total frequency, may be found from the probability table corresponding to the values of a_1/r and a_2/r. This fact is illustrated in Fig. 4(*b*) in which a_1/r is assumed as 1.0; the corresponding value of total frequency taken from the probability table is 0.5000 ($= P_1$). Similarly, with $a_2/r = 2.0$, $P_2 = 0.8226$, etc.

The duration curve of Fig. 4(*b*) is plotted as a full line at natural scale. It can be reduced to a straight line by a suitable change in the spacing of the percentage-of-time scale. Thus, let X_1 = scale value of P_1, in inches on the "probability paper"; and let X_2 = scale value of P_2, in inches. If X_1/X_2 is made equal to a_1/a_2, or to $\dfrac{a_1/r}{a_2/r}$ (as in the probability table), the duration curve will plot in a straight line.

TABLE 3.—Computation of Probability Scale

(Assuming Distance from 0.1% to 99.9% As 7.62 In.; from 50% to Either 0.1% or 99.9% As 3.81 In.)

Percentage of time		Relative distance from 50% line	P_a	$X = 0.4796\,a/r$	a/r	Distance from 50% line (in.)
50		0	0	0	0	0
40	or 60	0.10	0.20	0.1791	0.3756	0.31
30	70	0.20	0.40	0.3708	0.7775	0.65
25	75	0.25	0.50	0.4769	1.0000	0.83
20	80	0.30	0.60	0.5951	1.2479	1.04
10	90	0.40	0.80	0.9062	1.9001	1.58
8.87	91.13	0.4113	0.8226	0.9538	2.0000	1.66
5.0	95.0	0.45	0.90	1.1631	2.4389	2.03
2.15	97.85	0.4785	0.9570	1.4307	3.0000	2.49
2.0	98.0	0.48	0.96	1.4522	3.0451	2.53
1.0	99.0	0.49	0.98	1.6449	3.4490	2.87
0.5	99.5	0.495	0.99	1.8213	3.8190	3.18
0.1	99.9	0.499	0.998	2.1850	4.5817	3.81

For example, assume that the distance on the plotting paper between 0.1% and 99.9% is to be d inches, and it is desired to determine the proper scale distance between 25% and 75%: $P_{a1} = 75\% - 25\%$, or 0.50 for which

[28] "A Short Table of Integrals," by Benjamin Osgood Peirce, Ginn & Co., New York, N. Y., 1910.

[29] "American Civil Engineers' Handbook," Thaddeus Merriman, Editor, John Wiley & Sons, Inc., New York, N. Y., 5th Ed., 1930, p. 79.

$a/r = 1.00$; and $P_{a2} = 99.9\% - 0.1\% = 0.998$, for which $a/r = 4.5817$. The scale distance from 25% to 75% will then equal $(1.00/4.5817)d$ in. Computation of other plotting positions on the probability paper is shown in Table 3.

The values of X, corresponding to P_a are taken from a published table[28] of probability integrals.

Percentage of Time for Plotting Variability Index.—When a straight line is plotted on arithmetical probability paper (arithmetical scale for ordinates),

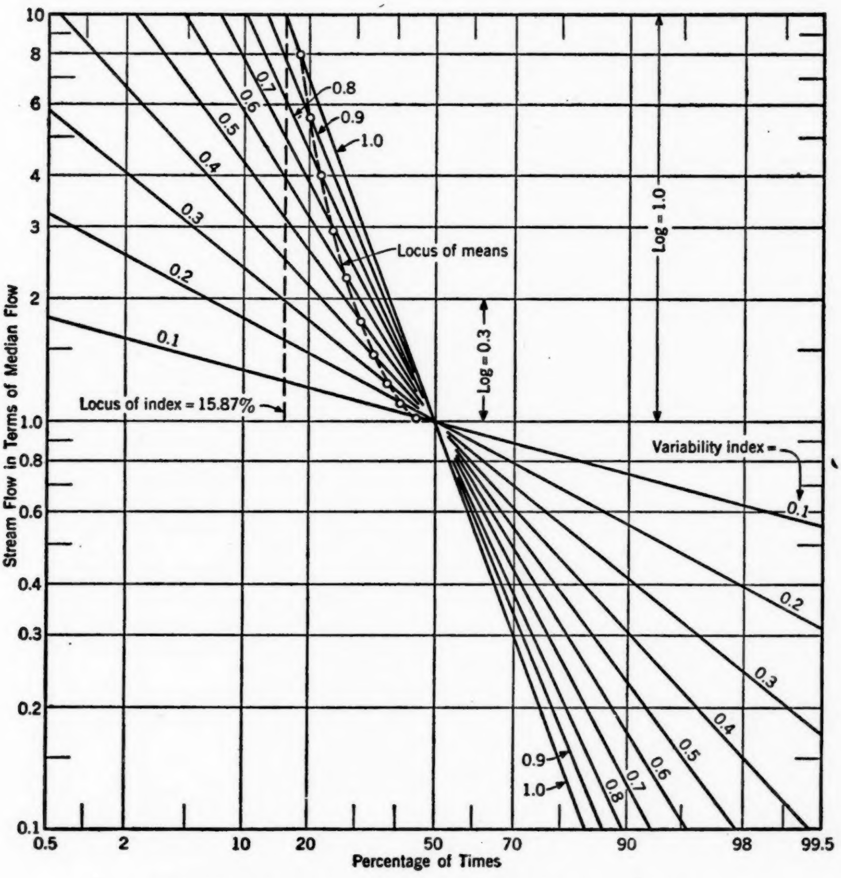

Fig. 5.—Daily-Flow-Duration Curves; Log Probability Scale

the difference between the ordinate value at any given percentage of time and that for 50% of time will be proportional to the probability integral, P_a, corresponding to the assumed percentage. Similarly, if a logarithmic scale is used for the ordinates, the corresponding difference between the logarithms will be proportional to P_a. The problem is to determine the percentage of time at which the logarithmic ordinate of the assumed straight line has a numerical value equal to the variability index of the line. This could be done

by scaling off the ordinates on the logarithmic probability paper at the 5%, 15%, ⋯ 95% points, and computing the index as explained by the authors; but a more accurate result will be obtained by assuming logarithmic ordinates proportional to the corresponding values of the probability integral, computing the corresponding values of the index, and then determining the percentage of time or probability (P_a) that would have a value of a/r equal to the index.

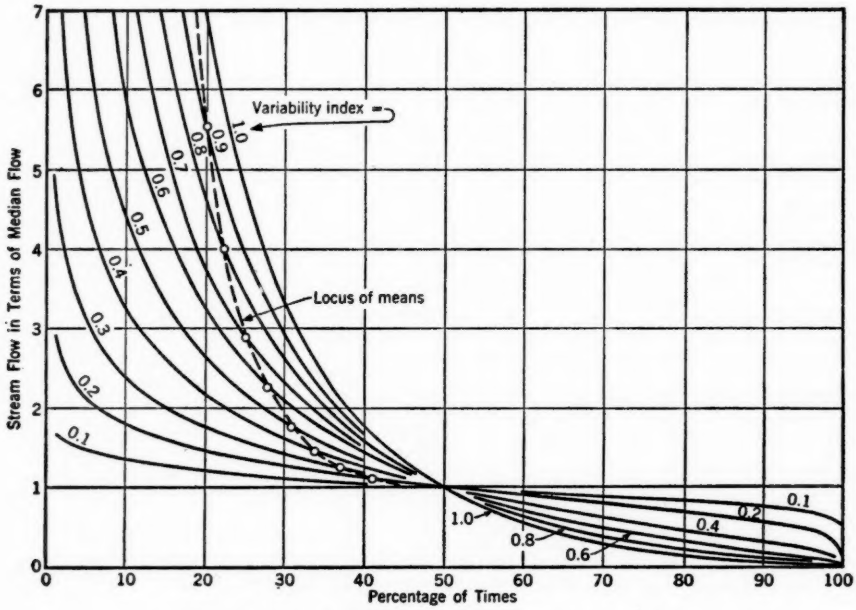

FIG. 6.—DAILY-FLOW-DURATION CURVES; NATURAL SCALE

The computation is shown in Table 4, in which Col. 1 denotes the range in percentage of time ($= 50\% \pm P_a/2$), and the data in Col. 3 are selected from available probability tables.[28] The variability index is computed with $\Sigma(a/r)^2$ in Col. 5, thus: The variability index from Table 4, is $\sqrt{\frac{\Sigma(a/r)^2}{n-1}} = \sqrt{\frac{19.3422}{9}}$ $= 1.4660$.

TABLE 4.—COMPUTATION OF VARIABILITY INDEX

Percentage of time (1)	P_a (2)	$X = 0.4769 \frac{a}{r}$ (3)	$\frac{a}{r}$ (4)	$\left(\frac{a}{r}\right)^2$ (5)
5 to 95	0.90	1.1631	2.4389	5.9482
15 to 85	0.70	0.7329	1.5368	2.3617
25 to 75	0.50	0.4769	1.0000	1.0000
35 to 65	0.30	0.2725	0.5714	0.3265
45 to 55	0.10	0.0889	0.1864	0.0347
Sum.................	....	....	....	9.6711
Doubled..............	....	....	....	19.3422

The corresponding value of $X = 1.4660$ times $0.4769 = 0.6991$; and P_a, corresponding to $X = 0.6991$, is 0.67718. Then, the percentage of time, for the computed index, is $50 - \frac{67.718}{2} = 16.14\%$. The discrepancy between

TABLE 5.—CHECK COMPUTATION OF VARIABILITY INDEX

Percentage of time	Q	Log Q	(Log $Q)^2$
5	6.65	0.8228	0.67700
15	3.30	0.5185	0.26884
25	2.17	0.3365	0.11323
35	1.56	0.1931	0.03728
45	1.165	0.0663	0.00440
$\Sigma(\log Q)^2$	....	....	1.10075
Doubled	....	....	2.20150

this value and 15.87% as given by the authors may be due to the omission of the decimals.

A series of duration curves with values of the variability index from 0.1 to 1.0 is shown in Fig. 5. For convenience in plotting, the 50% or median stream flow value is assumed as 1.0 (logarithm = 0.0). For an index of 1.0, the line will pass through a value of 10.0 (logarithm = 1.0) at 15.87% of time. The other lines will have arithmetical ordinates, at the same percentage of time, equal to their respective index values.

To show the true relative shape of these duration curves, they have been replotted in Fig. 6 on arithmetical scales. A check of the accuracy of this method is given in Table 5, in which the values of flow (Q) were scaled from the line plotted for the index = 0.50. As in the case of Table 4, the variability index is $\sqrt{\frac{2.20150}{9}} = 0.495$.

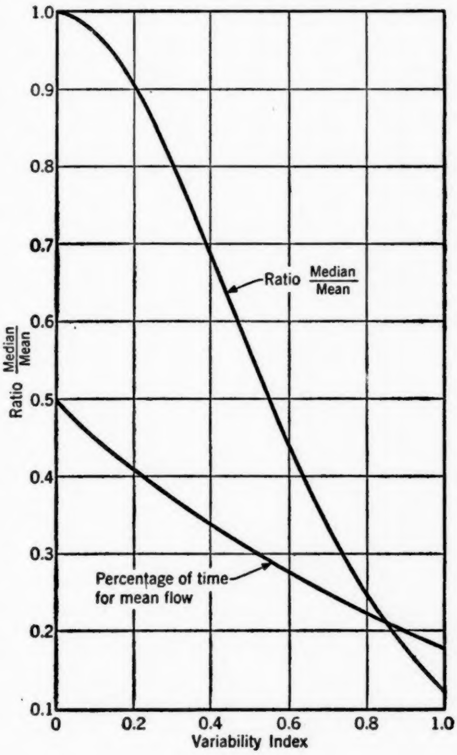

FIG. 7

Ratio of Median to Mean Value of Stream Flow.—Unless the duration curve of stream flow has been actually calculated, there is no method of determining the median value from the original records, although the mean stream flow is generally known or may be estimated. The method of constructing the duration curve proposed by the authors requires

the use of the median, and consequently it is necessary to establish a relation between these two values for various values of the index. The writer has attempted to establish such a relation by arithmetical integration of the plotted curves. The results are shown in Fig. 7 which also indicates the percentage of time at which the duration curve crosses the mean ordinate.

C. S. Ospina[30] and G. Tama,[31] Jun. ASCE.—The "variability index" offers an excellent method for the hydrologist to compare conditions in different watersheds. Its computation is simple and it does quite accurately represent the greater part of the flow variations. To facilitate plotting adjusted curves defined by the mean flow and the variability index, the writers suggest the construction of a graph which shows the percentage of time during which the mean flow is exceeded, as a function of the variability index (Fig. 8). For indexes

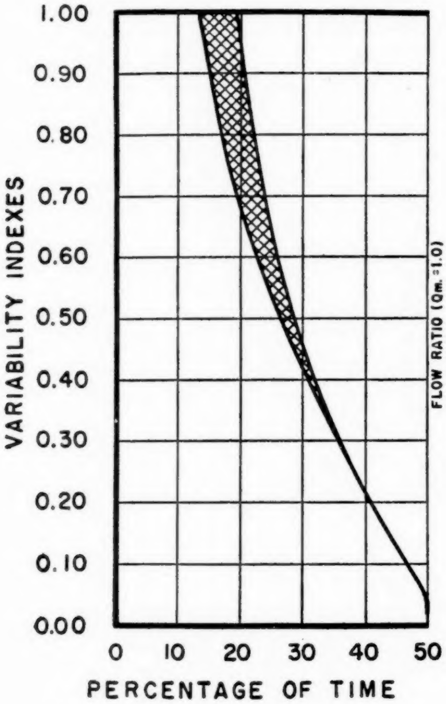

Fig. 8.—Adjusted Unit Duration Curves; Percentage of Time During Which Mean Flow Is Exceeded

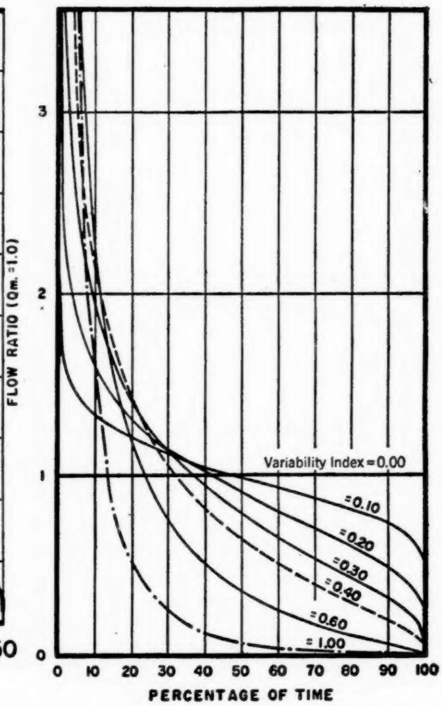

Fig. 9.—Adjusted Duration Curves for Different Variability Indexes

greater than 0.35 only a range can be indicated, with sufficient precision for practical purposes, because the average is affected by high discharges occurring during a small percentage of the time. Fig. 9 shows the adjusted duration curves used, which are straight lines on logarithmic probability paper.

---

[30] Partner, Olarte, Ospina, Arias and Payán, Engrs., Cali, Colombia.

[31] Associate, Olarte, Ospina, Arias and Payán, Engrs., Cali, Colombia.

The published data reveal no definite relation between mean unit flow (cubic feet per square mile) and the variability index, except for a certain similarity in rivers situated in the same part of the United States. It is noted that more than half of the 224 stations analyzed have drainage areas of less than 1,000 sq miles, and that the authors' conclusions in regard to the variability index incline to refer to small rivers. Both common sense and study of the tables confirm a definite influence of the size of the watershed on the index. The values in Table 6 have been deduced from the data presented by the authors. It is apparent that extremely high indexes occur only in small watersheds and that there is a general tendency toward lower values as the drainage area increases.

TABLE 6.—VARIABILITY INDEXES

Drainage area (square miles)	No. of stations	Most frequent	Maximum	Minimum
0.1 to 1,000.....	127	0.40 to 0.50	1.17	0.14
1,000 to 5,000...	59	0.40 to 0.50	0.96	0.18
5,000+........	37	0.30 to 0.40	0.79	0.025

For the Cauca basin in Western Colombia the following information can be given: Flow variations on the main river and most of the tributaries follow the same trend, regardless of the size of the watershed. An index of 0.23 is typical. The mean annual runoff is exceeded during 40% of the time and it averages 1.8 cu ft per sec per sq mile. In contrast to the uniformity of the variability index the unit runoff varies considerably even in watersheds with similar rainfall characteristics. The total drainage area at the lower end of the so-called Cauca Valley is 9,000 sq miles, the greater part of which is located in rugged country. Soils vary from impervious silts to pervious alluvial deposits in the flat regions; in most of the mountainous regions the soils are impervious with grass, bush, and forest cover. There are two dry seasons and two rainy seasons and the annual rainfall is about 40 in.

After years of work in irrigation and power planning, the writers were surprised to find a variability index so low in comparison to that for streams in the United States.

The authors recommend their procedure for studies where the available hydrological records are meager and they rightly state some of the factors which limit its use. The annual mean flow is to be determined from short-term flow data, from comparison with other hydrological records and drainage basins, from temperature records, etc. The selection of the variability index is to be governed by geological conditions and surface storage only, whereas rainfall distribution and intensity, vegetation cover, and arrangement of tributaries are acknowledged to be of influence to an unknown degree.

Actually, the number of factors is such that no accurate predictions can be made unless flow records exist covering a certain period. To the writers' knowledge, even short-term discharge data are likely to be more useful than coefficients derived from conditions in other regions. Attention is called to the fact that geologic conditions may be defined for restricted areas, yet in large basins the problem will become rather complex. Preliminary estimates of stream flow in undeveloped countries are hampered by deficient hydrological records and by little or no information as to the area and general features of

the drainage basin. The first guess will reflect personal judgment, whatever estimating method is used. As investigations proceed, flow records from newly established gaging stations will provide more reliable information. Older hydrological records, if existent, may be used to establish correlation. Unit duration curves, which are based on mean flow and have the advantage of being dimensionless, as well as chronological records from other rivers, are valuable to determine flow variations (variability index), annual mean discharge, and duration of droughts on the river studied. Mass curves of rainfall will be found useful to analyze the general trend. The accuracy of stream flow records should be checked, since systematic errors in high discharges, due to deficient section rating, will affect the variability index.

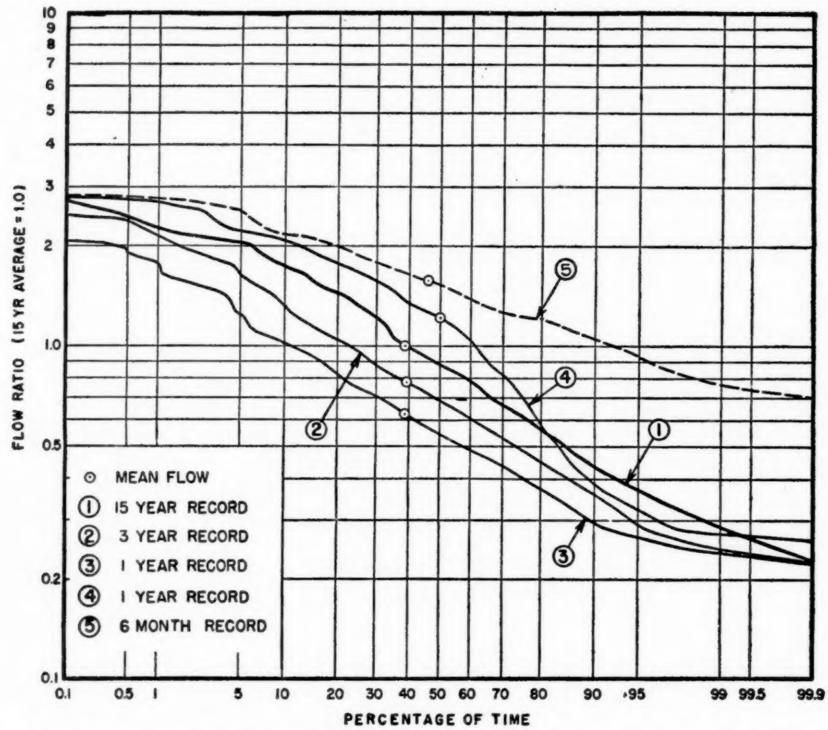

FIG. 10.—VARIATION OF DURATION CURVES FOR DIFFERENT PERIODS ON THE CAUCA RIVER AT CALI, COLOMBIA

The adjustment of short-term records to fit average conditions requires experience, since both mean flow and variability index change with the length of the period considered. Fig. 10 shows the actual duration curves of the same river (Cauca) for different periods, related to the 15-year average of $Q_M = 1$. The mean flow of the other periods is indicated by circles. It is evident that the "index" method of adjusting duration curves to straight lines on logarithmic probability paper should be applied only to long-term records. It

is necessary to examine thoroughly all pertinent relations until a representative duration curve can be plotted.

Finally, it should be remembered that most hydraulic studies will require an estimate of critical conditions during a period of severe drought. These are not fully defined by a duration curve representing a long-term record, and they have to be determined by a design hydrograph based on actual discharge data and adjusted to the most critical conditions.

The writers' experience on the hydrological studies in the Cauca Valley shows that, when only short-term records are available, the most difficult task is the determination of the minimum flow hydrograph—which represents minimum flow itself but also the hydrograph for the entire critical dry season. This hydrograph is of the greatest importance in planning irrigation and power projects where good storage reservoir sites are not available. The ratio of minimum flow to average flow for a period of from 3 to 15 years varies from 1:2.3 to 1:5.8 for the tributaries of the Cauca River, all with practically equal variability indexes. The Cauca itself has a ratio of 1:3.6. On the other hand, the characteristics of the dry season hydrograph vary very widely from one river to the other.

All this indicates the complexities and uncertainties of hydrological extrapolations. The writers consider that the greatest usefulness of the variability index, where few records are available, is to help develop the judgment of the engineer and thus aid him to make a better guess of the mean and the minimum flows.

The principal value of the paper is considered to be the development of a useful statistical investigating procedure and the analysis of a considerable volume of flow data from North American rivers. It is then hoped that other hydrologists will contribute information from different regions and nations in order to further the study of the factors which affect the index.

Raphael G. Kazmann,[32] Assoc. M. ASCE.—The authors would do a great service if they would enlarge on the use of their variability index. For example, the method depends on the proper selection of the variability index as outlined in Table 2. This table translates qualitative descriptions of geology, hydrology, and topography into quantitative terms. Apparently a rather thorough geologic and topographic survey of the drainage area is needed, together with a considerable amount of soil sampling and testing if comparable results are to be obtained by independent observers.

What permeabilities and what percentages of the drainage area are meant by the various subheadings under "bedrock and soil?" What slopes and percentages of drainage area are referred to under "relief?" Likewise what, quantitatively, is the average permeability of moraines? Under "lakes and swamps" what percentages of drainage area correspond to "small," "moderate," and "large?"

Then, there is the problem of determining the mean annual runoff of the stream. Using the method given by Don Johnstone, Assoc. M. ASCE, and

---

[32] Hydrologic Engr., Ranney Method Water Supplies, Columbus, Ohio.

William Perry Cross,[33] M. ASCE, the length of stream flow record needed to estimate the mean annual flow to various degrees of accuracy, for three Ohio streams, is shown in Table 7. These streams have been gaged continuously

TABLE 7.—Length of Record Needed

River	Location	Drainage area (sq miles)	Length of Record, in Years Required to Estimate Mean Annual Flow Within:			
			10%	15%	20%	25%
Licking...	Toboso, Ohio	672	54	25	15	10
Scioto....	Chillicothe, Ohio	3,847	62	29	17	12
Mad......	Springfield, Ohio	485	53	25	15	10

since 1923. On an ungaged stream, how many years of gaging do the authors suggest, and what degree of accuracy is desired, in determining the mean annual runoff for use with Table 2?

The writer would like to withhold judgment on the proposed method of developing flow duration curves until studies are made on watersheds whose physical characteristics are known and where long-term gaging records are available. The magnitude of deviation between the computed duration curves and the actual duration curves would show conclusively how applicable Table 2 is to field conditions. Naturally, all the suggested studies would have to be made without reference to actual long-term gaging records if valid comparisons are to result.

Jack Bruin[34] and H. E. Hudson, Jr.,[35] M. ASCE.—The proposed method of evaluating stream flow variability is obviously useful in applying data from streams with long periods of record to comparable streams for which short periods of record exist.

There have been a number of developments of small Illinois watersheds for municipal supply purposes. It is expected that there will be an increase in the use of smaller watersheds for impounded supplies to minimize sediment damage and spillway costs. Periods of zero flow exceeding 5% of the time are often encountered in such streams, making the method of this paper inapplicable. Similar situations are found in streams flowing on highly permeable valley fills and in arid regions where rainfall is infrequent. The method is worthy of modification to make it applicable to intermittent streams.

An attempt was made to compare the stream flow in Money Creek above Lake Bloomington with other gage records in central Illinois. For this purpose stream flow variability indexes were computed. An immediate difficulty made itself evident in that Money Creek had a zero flow 13.54% of the 12-year period of record. The variability index became infinity.

An inspection was made of the stream flow records from Illinois gaging stations, and it was found that this problem existed for five watersheds in the

[33] "Elements of Applied Hydrology," by Don Johnstone and William Perry Cross, Ronald Press Co., New York, N. Y., 1948.

[34] Eng. Asst., Eng. Sub-Div., State Water Survey Div., Dept. of Registration and Education, Urbana, Ill.

[35] Head, Eng. Sub-Div., State Water Survey Div., Dept. of Registration and Education, Urbana, Ill.

group from 45 sq miles to 500 sq miles. The largest of the gaged areas having zero flow more than 5% of the time is Mazon River, which drains a watershed only a small part of which is morainic, the remainder being largely lacustrine.[36] Money Creek drains a watershed of 45 sq miles, which lies between two moraines of Wisconsin age. The following approximate distribution of subsoil permeability[37] exists in the two watersheds (areas in square miles):

Permeability	Mazon River	Money Creek
Slowly to very slowly	173	0
Moderately slow	212	7
Moderate	85	38
Total	470	45

The Lake Bloomington watershed is somewhat representative of the average subsoil permeability in Illinois, whereas the Mazon River subsoils are considerably less permeable than most subsoils in the state. Zero flow was reported for Mazon River during 7.06% of the period of record, which is 9 years. In this period the run-off from streams in the same region was slightly above normal. Station discharge duration curves for Money Creek above Lake Bloomington and Mazon River at Coal City, Ill., computed from records gathered through cooperation with the United States Geological Survey, are shown in Fig. 11.

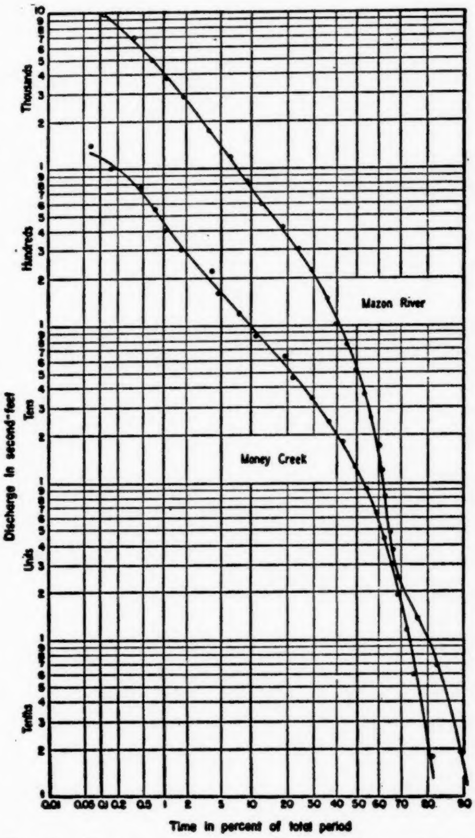

Fig. 11.—Discharge Duration Curve, Money Creek and Mazon River in Illinois

In view of the inapplicability of the authors' method to streams in which there is zero flow more than 5% of the time, it is suggested that the variability index be based on the period during which flow exists. Calculated values for the streams discussed herein, using the percentage values and procedure proposed by the authors are as follows:

[36] "Physiographic Divisions of Illinois," by M. M. Leighton, George E. Ekblaw, and Leland Horberg, *The Journal of Geology*, January, 1948, p. 16.

[37] "Principal Soil Association Areas of Illinois," Dept. of Agronomy, Univ. of Illinois Agri. Experiment Station, Urbana, Ill., May, 1948.

Description	Mazon River	Money Creek
Variability index	1.27	0.94
Percentage of time of zero flow	7.06	13.54

This index would be accompanied by a statement of the proportion of time during which there is zero flow.

H. W. Lull.[38]—A particularly interesting feature of this ingenious index of variability, not pointed out by the authors, is its comparability in meaning to a coefficient of variation—that is, it represents not variation, as such, but variation in respect to the mean. Since a constant logarithmic deviation corresponds to a constant percentage change, computing the standard deviation of the logarithms of the discharge gives an index of deviations in percentage of the mean, as is done directly by the coefficient of variation.

An undesirable feature of the index is that it cannot be related directly to mean flow. However, because of the normality of the distribution of discharge, expressed in logarithms, throughout most of the range, the index when used with the logarithm of mean flow provides estimates of discharge at various frequencies. No estimates, obviously, could be made for the curvilinear parts of the log discharge frequency relationship. In this connection, perhaps the utility of the index would be increased if estimates of discharge at the 1% and 99% frequencies were appended to it. These values, by indicating the nature of the curvilinearity of the extremes, would be of importance to those interested either in flood or in minimum flows.

An interesting point is the relation of the variability indexes to the drainage area. The authors state that streams from large watersheds will tend to have lower indexes than streams from small watersheds because of less synchronization of flow and greater channel storage. Surprisingly, a tabulation of the area and the variation of 22 streams, listed by the authors (each of which were gaged at 2 or more stations), reveals that this is not the case (see Table 8). Seemingly, the indexes should decrease in the downstream direction. According to the tabulation, this occurred on 11 streams, whereas on 8 streams the indexes became larger, and on 3 streams (Table 8(a)) they remained constant. Closely associated with a decrease in the variability indexes is an increase in the mean flow; and with an increase in the indexes, there is a decrease in the mean flow (see Cols. 4, Table 8). Since the indexes reflect variation in respect to the mean, these relationships permit the speculation that either variation remains constant, or, more likely, that it changes in the same direction but less rapidly than does mean flow.

After considering several factors which affect stream flow variability, the authors concluded that geology and soils, and surface storage had the major influence. Although there is some evidence to support this conclusion, it is hardly sufficient to warrant the finality the authors give to it. In a study of this kind, it is extremely doubtful if factors having an important bearing on stream flow variability can be isolated.

One reason for this doubt is the complication posed by the wide range of drainage areas. On small drainages, vegetation and weather may be important

[38] Forester, Div. of Forest Influences, Forest Service, U. S. Dept. of Agriculture, Washington, D. C.

factors, local occurrence of high intensity rainfall or droughts, or density of vegetation, having a greater effect on variability than they would have on larger areas. Geology and soils could well be the most influential factor on larger areas possessing a variety of different types of vegetation. Similarly, a

TABLE 8.—RELATION OF STREAM FLOW VARIATION TO SIZE OF DRAINAGE AREA

Station[a] (1)	Drainage area[b] (2)	Index[c] (3)	Mean flow[d] (4)
(a) VARIATION REMAINS CONSTANT WITH INCREASE IN AREA			
Muskingum River, Ohio (Table 1(*d*))			
36	5,982	0.48	1.021
37	7,411	0.48	0.791
Lamoille River, Vermont (Table 1(*g*))			
4	250	0.38	1.73
3	310	0.38	1.64
Winooski River, Vermont (Table 1(*g*))			
7	338	0.43	1.43
6	1,018	0.43	1.66
(b) VARIATION DECREASES WITH INCREASE IN AREA			
Des Moines River, Iowa (Table 1(*a*))			
3	5,490	0.64	0.278
5	6,818	0.63	0.267
6	13,200	0.59	0.246
7	13,900	0.50	0.382
Skunk River, Iowa (Table 1(*a*))			
9	320	0.94	0.350
10	2,890	0.53	0.495
11	4,290	0.54	0.502
Iowa River, Iowa (Table 1(*a*))			
12	1,500	0.51	0.471
13	3,230	0.49	0.457
14	12,480	0.38	0.489
Cedar River, Iowa (Table 1(*a*))			
15	845	0.39	0.341
18	1,660	0.38	0.407
19	6,640	0.36	0.442
(b) VARIATION DECREASES WITH INCREASE IN AREA (*Cont'd*)			
Wapsipinicon River, Iowa (Table 1(*a*))			
20	1,060	0.64	0.442
21	2,300	0.45	0.564
Sandusky River, Ohio (Table 1(*d*))			
12	299	0.75	0.833
2	1,248	0.63	0.702
Scioto River, Ohio (Table 1(*d*))			
18	988	0.69	0.784
32	1,624	0.62	0.843
43	3,847	0.55	0.889
Miami River, Ohio (Table 1(*d*))			
27	1,155	0.54	0.846
28	2,513	0.44	0.921
39	3,639	0.46	0.966
Ashuelot River, New Hampshire (Table 1(*h*))			
8	71.1	0.50	....
9	420	0.45	....
Connecticut River, Massachusetts (Table 1(*i*))			
6	8,000	0.39	....
5	8,390	0.36	1.454
Delaware River, New York and New Jersey (Table 1(*l*))			
9	875	0.47	2.331
7	3,070	0.44	1.825
10	6,800	0.35	1.690
(c) VARIATION INCREASES WITH INCREASE IN AREA			
St. Croix River, Wisconsin (Table 1(*b*))			
1	1,550	0.19	0.761
2	2,820	0.24	0.705
4	5,120	0.31	0.583
5	5,930	0.30	0.566
Wolf River, Wisconsin (Table 1(*b*))			
18	812	0.20	0.958
29	2,240	0.28	0.820
Kankakee River, Illinois (Table 1(*c*))			
7	2,340	0.32	0.762
6	4,870	0.42	0.640
Sangamon River, Illinois (Table 1(*c*))			
11	550	0.76	0.763
10	2,560	0.77	0.696
Maumee River, Ohio (Table 1(*d*))			
7	2,049	0.56	0.744
1	6,314	0.79	0.671
Auglaize River, Ohio (Table 1(*d*))			
9	333	0.56	0.883
8	2,329	0.73	0.696
Tuscarawas River, Ohio (Table 1(*d*))			
23	1,398	0.43	1.000
22	2,439	0.47	1.007
Mad River, Ohio (Table 1(*d*))			
30	485	0.29	1.053
29	632	0.30	1.050

[a] Cols. 1, 2, and 3, Table 1. [b] Col. 4, Table 1. [c] Col. 6, Table 1. [d] Col. 7, Table 1.

small area possessing a uniformity of topography would produce stream flow of an entirely different character than larger drainages including a diversity of topographical features. Thus, it is difficult to select factors of major importance when watersheds of different sizes are considered.

Furthermore, unless each of the several factors is given equal consideration, it is doubtful if any conclusion is justified as to relative importance. At the present time, lack of data on all factors would prohibit such an analysis; but it may be that if the authors had given more attention to factors considered of minor influence, a different conclusion would have been reached as to relative importance of the factors.

For instance, from the authors' conclusions the inference can be made that vegetation is not an influential factor. A comparison of the variation of streams

from forested and nonforested areas of similar size in Iowa and Wisconsin does not support this view (see Table 9). Furthermore, it should be emphasized that the relatively small variation from areas of greatest relief, in Iowa and North Carolina, is perhaps due more to the fact that they are mainly forest covered than to the reasons given by the authors. If weather were truly considered a factor, perhaps the authors should have considered the effect of the accumulation and rapid melting of snow on the variability in the more northern

TABLE 9.—COMPARISON OF VARIABILITY INDEXES OF FORESTED AND NONFORESTED DRAINAGES

FORESTED AREAS (SQ MILES)			NONFORESTED AREAS (SQ MILES)		
River	Area	Index	River	Area	Index
(a) IOWA					
Little Maquoketa	130	0.39	Skunk	320	0.94
Yellow	224	0.35	Lime Creek	535	0.63
Turkey	1,530	0.44	Iowa	1,500	0.51
			Shellrock	1,700	0.46
			Cedar	1,660	0.38
			Wapsipinicon	1,060	0.64
(b) WISCONSIN					
Apple	550	0.24	Crawfish	732	0.57
Jump	510	0.62	Rock	971	0.68
Pine	543	0.32	Milwaukee	661	0.48
Eau Claire	500	0.53	Sugar	529	0.24
Black	756	0.60			
Wolf	812	0.20			
Oconto	678	0.26			
Peshtigo	571	0.36			
Mean variability (Table (a))		0.39	Mean variability (Table (a))		0.49

states. In addition, the factors deemed important by the authors, when considered singly and in respect to the area involved, are not of major importance. Surface storage can obviously be an important factor only in the relatively limited areas in which it exists. Geology and soils were found to be important in certain areas, but not in enough areas to justify selection as a major factor.

Perhaps a more reasonable conclusion would be that the relative importance of the factors cannot be determined for as large an area as the one concerned. All the factors may be considered as having a bearing on variability; each is present to a degree in all drainages; in certain drainages one or two factors may be of particular importance because they are present in sufficient intensity to mask the effect of the other factors. However, because of the variability of the factors over the particular area, in addition to the variation in size of the drainages which tends to intensity the effect of the more variable factors on the smaller areas, and the fact that evidence is not available to permit a thorough analysis of the effect of each factor, the data that the authors present do not allow an evaluation of the relative importance of the factors affecting stream flow variability.

WILLIAM D. MITCHELL,[39] M. ASCE.—The authors have made a sound and comprehensive presentation of stream flow duration. Their use of logarithmic probability scales is particularly commendable.

Instances are rare indeed in which the logarithmic probability plot of the flow duration curve will yield a straight line throughout the entire range of flow. For such rare cases, the curve is adequately described by the two parameters suggested by the authors—namely, the variability index and the mean discharge. As they have pointed out, the position of the curve must be established by trial, but that task is not particularly difficult. On the other hand, in all cases with which the writer is familiar, there is substantial curvature near one or both ends of the logarithmic probability plot so that the two parameters are not adequate to define the curve. This fact has been recognized by the authors, as evidenced by their statement (under the heading, "Use of the Variability Index"): "The ends of this line for a short distance on each end will have to be altered to match the best possible estimates of maximum and minimum flow * * *."

A project on which the writer is engaged involves the flow duration curves for 15 of the Illinois stations used by the authors. The basic data differ in only one major respect—namely, the records have been extended to include all complete water years through 1945. One of the stations, Big Muddy River at Plumfield, Ill., has been selected to show the relation between the synthetic curve obtained by the authors' method, and the actual curve derived from the station data.

In the illustration the synthetic development assumed that the variability index and the mean discharge per square mile (q) were known to be, as given in Table 1, Cols. 6 and 7, 1.01 and 1.021, respectively. From published data,[40] it was determined that the extremes of discharge prior to September 30, 1933, were 16,300 cu ft per sec and 0 cu ft per sec (the equivalent of 21.6 cu ft per sec per sq mile and 0 cu ft per sec per sq mile) and it was inferred that the time of no flow was about 9 days in 19 years, or a little more than 0.1%.

The resulting synthetic curve appears as the continuous line in Fig. 12. In determining the slope of this line, it was necessary to modify the statement of the authors (under the heading, "Use of the Variability Index"):

> "The shape of the duration curve can be obtained by drawing a straight-line duration curve on logarithmic probability paper with a slope such that the ratio of the discharge exceeded 15.87% of the time to the discharge exceeded 50% of the time was equal to the variability index selected."

Since the selected index is 1.01, it appears that a line of such slope would be nearly horizontal, and not at all representative of the flow duration curve. Obviously, the ratio of the discharges should be equal to the antilogarithm of the selected index, or, in this case, 10.23. The line was drawn on this slope and, on the first trial, was made to pass through the point given by J. H. Morgan[41]—q = 1.01 and percentage of time equals 25.8. The data on extreme

---

[39] Hydr. Engr., U. S. Geological Survey, Champaign, Ill.

[40] "Surface Water Supply of the United States, 1933," *Water-Supply Paper No. 745*, U.S.G.S., Washington, D. C., 1935, p. 213.

[41] "Flow-Duration Characteristics of Illinois Streams," by J. H. Morgan, *Transactions*, Am. Geophysical Union, Vol. 17, 1936, p. 419.

discharges offered no reason for curvature in the lower end of the line, but the upper end obviously must become asymptotic to the horizontal $q = 21.6$, and the line was drawn accordingly. Step integration of the area under this line yielded a mean discharge of 1.25 cu ft per sec per sq mile, greatly in excess of the assumed value of 1.02 cu ft per sec per sq mile, which indicated that the curve must be moved leftward to the position as shown.

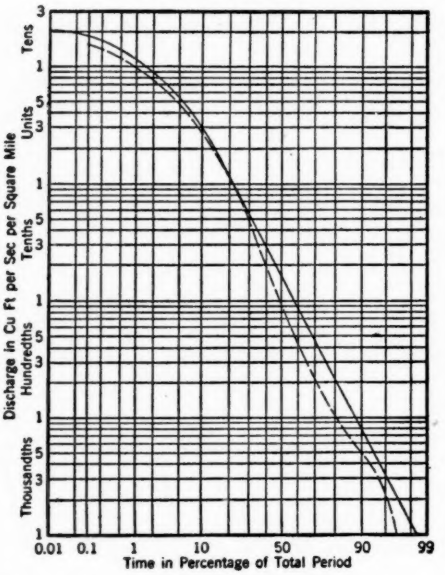

Fig. 12.—Duration Curves for Daily Flows of Big Muddy River at Plumfield, Ill.

For comparison, the duration curve as developed by the writer, based on actual records for the 35 water years, from 1909 to 1912 and from 1915 to 1945, is shown as the dashed line in Fig. 12. As indicated in Table 10, the variability index for the longer period is 1.08, whereas the mean discharge is 0.891 cu ft per sec per sq mile. The duration of the mean discharge is 23.5% of the total time.

The decrease in mean discharge, due to the use of additional record, is largely responsible for the downward shift from the synthetic to the actual curve. On a percentage basis this shift is a maximum at about the 70% duration point, where the synthetic curve indicates a discharge of more than double that of the actual curve. Had this particular point been the subject of investigation, the use of the synthetic curve could not be recommended. At other points, particularly in the region of the mean discharge, the agreement between the two curves is very satisfactory. It should be noted that the assumed shape of the synthetic curve has been helpful in obtaining the good agreement in the region near the mean discharge. Since the parameters for the synthetic curve are based on the same data as used by Mr. Morgan, a properly shaped duration curve for the years from 1915 to 1933 would have passed through the point given by him; that is, $q = 1.01$ and percentage of time equals 25.8. However, since the shape of the curve was only approximated by the synthetic procedure, it was necessary to compensate for the inaccuracy by shifting the synthetic curve to pass through the point $q = 1.01$ and the percentage of time equals 21.8. It is believed that the improvement thus obtained is only coincidental; in other cases the shift necessary to compensate for inaccurate shape may be such as to move the synthetic curve away from the actual curve in the region of the mean discharge.

Furthermore, the curvature in the upper part of Fig. 12 is opposite to that mentioned by the authors (unde the heading, "Use of the Variability Index").

This is due to the fact that Big Muddy River is bordered by extensive flood plains which, by their storage effects, materially reduce the magnitude of the high discharges.

The foregoing example benefits, of course, from the fact that very pertinent values were assumed for the index and the mean discharge. Had it been necessary to estimate these values from data given in Table 2 or from records of near-by watersheds, any less appropriate estimates would have been reflected in the synthetic curve.

TABLE 10.—CHANGE IN VARIABILITY INDEX AND MEAN DISCHARGE AT SELECTED SITES, DUE TO ADDITION TO LENGTH OF RECORD; STREAMS IN ILLINOIS (SEE TABLE 1(*c*))

Station No.	River	Station	Variability Index			Mean Discharge		
			Lane-Lei	Mitchell	Ratio	Lane-Lei	Mitchell	Ratio
1	Pecatonica	Freeport	0.26	0.30	1.15	0.766	0.677	0.88
2	Fox	Algonquin	0.39	0.42	1.08	0.614	0.560	0.91
3	Rock	Lyndon (Como)	0.30	0.32	1.07	0.694	0.597	0.86
7	Kankakee	Momence	0.32	0.30	0.94	0.762	0.747	0.98
8	Spoon	Seville	0.59	0.63	1.07	0.620	0.615	0.99
9	La Moine	Ripley	0.65	0.71	1.09	0.604	0.583	0.97
10	Sangamon	Riverton	0.77	0.75	0.97	0.696	0.678	0.97
11	Sangamon	Monticello	0.76	0.76	1.00	0.763	0.727	0.95
12	Vermilion	Danville	0.73	0.70	0.96	0.566	0.691	1.22
13	Macoupin Creek	Kane	0.85	0.86	1.01	0.721	0.738	1.02
14	South Fork	Kincaid	0.89	0.94	1.06	0.824	0.878	1.07
15	Kaskaskia	Vandalia	0.70	0.74	1.06	0.753	0.717	0.95
18	Skillet Fork	Wayne City	1.09	1.10	1.01	0.940	0.813	0.86
19	Big Muddy	Plumfield	1.01	1.08	1.07	1.021	0.891	0.87
20	Cache	Forman	1.02	1.07	1.05	1.400	1.23	0.88

The writer has had no opportunity to test the data given in Table 2. Some information may be provided, however, on the variation of the index at a given point on a given stream, as a result of variation in the period of record. As stated, new duration curves have been prepared for 15 of the stations described by the authors. These new curves include data for all complete water years prior to 1946. New values of the mean discharge and of the variability index, computed by the authors' method, are shown in Table 10. At certain stations, the index value has changed appreciably, and, in general, the index value has increased while the mean discharge has decreased. These fifteen cases are too few to warrant firm conclusions as to trends; but they do raise a question as to whether an increase in the value of the variability index may not be expected with an increase in the length of record.

C. R. OWNBEY,[42] JUN. ASCE.—This paper should help to stimulate the interest of hydraulic engineers in the possibilities of a powerful tool for the evaluation of hydrologic data, which heretofore has not received the attention it deserves. With some notable exceptions, hydraulic engineers have not kept pace with, or taken proper advantage of, the developments in the mathematical science of statistics. The authors have introduced, to the analysis of stream

[42] Asst. Prof. of Civ. Eng., Univ. of Tennessee, Knoxville, Tenn.

flow data, one of the characterizing numbers used by statisticians as a measure of the nature of the distribution of data.

In introducing their variability index (see under the heading, "Previous Studies of Duration Curve Properties") they state that:

> "Many civil engineers are familiar with this from its application in triangulation and base line measurement, where the standard deviation is known as the 'root mean square error.'"

They might also have stated another analogy that is perhaps even more familiar to civil engineers—that of the mathematical and physical similarity to the moment of inertia and the radius of gyration of areas. In characterizing stream flow data, the engineer is interested in the spread, or dispersion, of values, rather than in the probable error; and this is essentially what the radius of gyration measures for the cross-sectional area of a column or other structural member.

Without having seen the original, more detailed paper filed in the Engineering Societies Library,[42a] the writer believes that this paper has suffered in condensation from the omission of background material. The explanation of the method of constructing a synthetic frequency curve from the variability index and the mean flow, in particular, is not clear.

It is wondered why the authors chose to depart from the accepted definition of "standard deviation" in defining their variability index. The defining formula for standard deviation is

$$\sigma = \sqrt{\frac{\Sigma(X - \bar{X})^2}{N}} \quad \text{(7)}$$

in which X is an individual variate; $\bar{X}$ is the mean value; and N is the total number of variates in the population group, which, in the case of daily stream flows, would mean the number of days from minus to plus eternity. When the calculation of standard deviation is based on a "sample" consisting of n days of record, a fair approximation is obtained by applying a correction factor $\sqrt{n/n - 1}$, making Eq. 7 equal to

$$\sigma = \sqrt{\frac{n}{n-1}} \sqrt{\frac{\Sigma(X - \bar{X})^2}{n}} \quad \text{(8)}$$

The precision of the number so obtained increases with the number of days, n, used in the calculation, and it will be noted that the correction factor approaches unity as n increases.

In effect, the authors have applied the correction factor in Eq. 8 by using the divisor 9 for their calculations based on the 10 values picked from a frequency curve, which is an overcorrection (as will be explained).

The use of 10 numbers picked from a graph is questionable. It would seem better to group the raw data into equal class intervals and to use midvalues of these intervals, with corresponding frequencies, for the calculation of the index. These intervals should be chosen with regard for proper statistical representation of the data, but would surely be such as to give more than 10 groups in

[42a] 29 W. 39th St., New York 18, N. Y.

this case. At the same time, the arithmetical simplifications made possible by the use of equally-spaced intervals would reduce the labor involved to a minimum. It is doubtful that the calculation (suitably made) of the index from the raw data would require any more time than a computation by the method of picking 10 points from the duration curve as described by the authors, especially in view of the time required to arrange and plot 3,650 values of daily flow for a 10-year period of record.

Referring again to the use of 9 as a divisor in calculating the "variability index," this would be the proper number to use if the index were based on only 10 values of raw data, except that, if the index had been based on only 10 days of record, it would be statistically worthless. The authors have stated that the minimum period used in the compilation of their tables was 10 years, which for daily flows means a minimum of 3,650 variates. Would it not have been more nearly correct to divide by 10, the number so obtained being theoretically subject to a correction factor equal to $\sqrt{3{,}650/3{,}649}$, which is negligible?

The principal value of probability paper in these studies would seem to be that it provides a ready, but rough, method of determining whether the data of the problems at hand follow a curve called the normal curve of distribution, with its related probability law—the proof being that points plotted on this paper yield a straight line. Apparently the authors concluded that the actual values of stream flow do not follow the normal law, but that the logarithms of these values do follow it—hence, the use of logarithms rather than actual stream flow values in calculating the index.

In any problem for which it can be assumed that the variates (whether raw data or their logarithms) follow the normal distribution curve, it is a simple matter to compute values for a frequency curve with percentage of time as its argument (called a flow-duration curve by hydraulic engineers). For this purpose it is necessary to have only the mean value and the standard deviation (known or estimated from tables such as those presented in this paper) and a table of solutions of the probability integral, readily found in references. However, one characteristic of the normal distribution is equality of the mean and the median. For stream flow it would be necessary to assume that the mean flow was the same as the flow equaled or exceeded 50% of the time, and this is not usually found to be true. Perhaps the authors' method of constructing a synthetic frequency curve is applicable to skew frequencies. If so, it is hoped that further explanation will be given in the closing discussion.

The numerical value of the standard deviation as applied to the flow of a stream reflects not only the variability of the flow but also its magnitude. The great value of data such as presented in this paper should be in the possible use of a stream as a "model," from which to construct or check a flow-duration curve for some comparable stream for which data are scarce. To eliminate the effect of size from such a comparison, it seems that the ratio of the standard deviation to the mean flow (the coefficient of variation) should be used, rather than the standard deviation itself. It should be understood, of course, that in order for 2 streams to be comparable, their sizes could not be greatly different —certainly not of the order of scale ratios commonly used for other purposes in hydraulics.

S. P. Wing,[43] M. ASCE.—In a different connection, John S. Cotton,[44] M. ASCE, emphasized the fact that the determination of reservoir capacity needs "* * * as much or more attention than most other current problems * * *" to which engineers habitually give their attention. The paper provides additional techniques through which such capacity studies can be made. Nearly all countries possess rainfall information from which reasonable estimates of mean annual runoff are possible; but knowledge of the variation of the daily stream flow that determines reservoir capacity is much more limited. This situation is particularly true in underdeveloped countries where knowledge of minimum stream flow is especially needed now because of current emphasis on wide-scale river basin planning. The paper deals with two problems: First, the development of an adequate index through which daily stream flow variation can be expressed; and, second, the determination of the factors from a knowledge of which its quantitative value can be estimated.

The utility of arithmetic averages for summarizing a mass of data is commonplace. Hydrologists guide their judgment through their knowledge of mean runoff per square mile, mean rainfall, mean flood discharges, and so forth. Knowledge of these mean values helps prevent mistakes when the engineer is dealing with unreasonable observations or fragmentary data. On the other hand, acceptable data on the mean values of stream flow variation are few, perhaps because of the lack of a generally accepted variation coefficient through which past studies could have been summarized. In placing before the profession their logarithmic "flow variability index," and in giving its value for some 200 American streams as computed from hundreds of thousands of observations, the authors provide hydrologists with type values for streams whose physical characteristics are well known and hence provide the engineer with a "judgment" background.

Long ago Francis Galton[45] pointed out that if, instead of characterizing the frequency distribution of sets of observations of natural phenomena such as runoff, rainfall, and so forth by standard deviations, they be characterized by the standard deviation of the logarithm of the observations, a much better and more useful variation index would be obtained, and one better fitting the observed distributions. Luigi Gherardelli[46] in Italy and Aime Coutagne[47] in France early in the 1930's utilized his suggestions in their studies of runoff variations. In a brilliant paper classifying streams not only by their indexes of variation with respect to daily stream flow as in the present paper, but also by a series of other variation indexes to take account of seasonal and subsoil seepage characteristics, Mr. Coutagne presents values for the Lane-Lei variation coefficient (in terms of its reciprocal) for a number of European streams. His values are very similar to those in the paper but his conclusions are different.

[43] Civ. Engr., Washington, D. C.

[44] Discussion by John S. Cotton of "Multiple-Purpose Reservoirs: A Symposium," *Proceedings*, ASCE, September, 1949, p. 1075.

[45] "Statistics by Intercomparison, with Remarks on the Law of Frequency of Error," by Francis Galton, *The London Edinburgh and Dublin Philosophical Magazine and Journal of Science*, 4th Series, Vol. XLIX, January to June, 1875, p. 33.

[46] "Au Sujet de Quelques Récentes Formules Statistiques Concernant les Déterminations Hydrologiques," by Luigi Gherardelli, *Rapports Bulletin No. 21*, Assn. Internationale d'Hydrologie Scientifique, Rome, Italy, 1934.

[47] "Classification des Cours D'Eau D'Après Les Méthodes Statistiques," by Aime Coutagne, Pithiviers, Imprimerie des Caisses d'Épargne, Grenoble, France, 1935.

Presumably Mr. Coutagne utilized standard methods similar to those presented by the American Society of Testing Materials[48] in computing the variation index, procedures which the writer would recommend. It is hoped, in their closure, the authors will clarify their own methods by illustrating with an example, as it is important that their method be fully understood by those utilizing their work and the present procedures are not clear.

A couple of illustrations of the physical meaning of the variation index should help expand its use. A stream whose logarithmic variability is 0.6 (the typical stream of Table 2) has an antilogarithm of 4. The theory of statistics indicates that a stream with this standard deviation of its logarithms of flow has two thirds of its daily flows within the range of one fourth to four times the yearly daily average. Average flow here means the flow corresponding to the antilogarithm of the average logarithm, a value of the order of 0.6 times the ordinary arithmetic mean flow. Such a stream for 83% of the time would develop 25% of its potential power without storage. The wide range in the suitability of the streams for power developments from a purely hydrological viewpoint is indicated by the wide range of the authors' indexes, from 1.2 to 0.2. These, having antilogarithms of 16 and 1.6, mean that for 83% of the time these streams had flows that were greater than 6% and 63% of the yearly average.

The second part of the paper concerns a discussion of the factors on which day-to-day stream flow variation depends. In the "Synopsis," it is stated:

> "* * * in the northeastern quarter of the United States the geology (including soil cover) and the presence of lakes and swamps are the most important factors in stream flow variability during the greater part of the time and * * * other factors have relatively smaller effect."

In view of the many varying characteristics included in the watersheds of the stream cited in the paper (such as rainfall, drainage area, geography, and culture—all of which are commonly thought to effect the uniformity of stream flow), it is regretted that the paper does not contain supporting analysis for this conclusion. In Table 2 the authors propose that the variation index for an unfamiliar stream be estimated by utilizing the average variation index of 0.60 typical of their data and then modifying it to values of from 0.7 to 0.2 (a ratio of 3.5 to 1) purely on the basis of effects due to varying geological factors—that is, percentages of glacial moraine and bed rock, or varying percentages of lakes or swamps. Other factors are given no consideration. The writer proposes briefly to examine whether other factors may not be of equal importance and demonstrated better.

Rainfall.—That varying rainfall causes a varying runoff is almost too well known to require comment. Correlations greater than 0.90 are commonly found when comparing these two variables whether on the basis of daily, weekly, monthly, or annual flows. If this is true a glance at typical rainfall variation indexes should give a guide to the variation to be expected of runoff. From Mr. Coutagne's data the following daily rainfall variation indexes are

---

[48] "Manual on Presentation of Data," Committee E-1, A.S.T.M., Philadelphia, Pa. April 1943.

presented for certain European stations:

Station	Index
Lyon, France	0.71
Marseille, France	0.57
Brest, France	0.55
The Alps	0.42

Rainfall at Lyon is 1.7 times as variable as at an Alps station, and, presumably, streams having similar drainage basins in the two locations would have runoffs varying in a similar manner. The writer suggests that, as more daily rainfall indexes become available from locations in the tropics to the thunderstorm belts of the arid west, the range of rainfall variation will be found quite comparable to the range in stream flow variation presented by the authors.

Altitude.—That varying altitude affects the runoff variability index is suggested by the authors in commenting on their North Carolina data. There they found that their index varied from 0.62 to 0.25 (a ratio of 2.5 to 1) as they

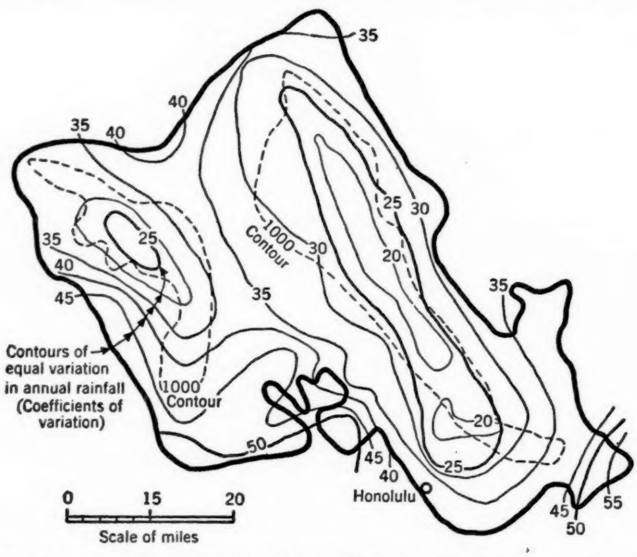

FIG. 13.—ANNUAL RAINFALL BECOMES MORE UNIFORM WITH INCREASE IN ELEVATION

went from the plains to the mountains and ascribed it to the more uniform mountain rainfall. That this effect exists is also shown in Fig. 13, from data supplied by Winters P. Nakamura.[49] Here in terms of the arithmetic deviation to which, roughly, the authors' index is proportional, the annual rainfall variation index changes from 50 to 20, a ratio of 2.5 to 1 as the altitude becomes greater.

Area.—The effect of the area of watershed on the uniformity of its drainage is likewise common knowledge. Big rivers tend to be much less variable than

[49] "Rainfall Coefficient of Variation," by Winters P. Nakamura, *Monthly Weather Review*, December 1933, pp. 354–360.

small flash streams. The authors' data, when plotted against drainage area, indicate that, as the size of the watershed increases from 100 sq miles to 10,000 sq miles, the variation index is reduced on the average from 0.76 to 0.36—a reduction to 2 to 1. Mr. Coutagne's data, shown as follows, indicate similar reductions, the ratio being 2.5 to 1:

Stream	Area (sq miles)	Coefficient
Reno, Italy	300	0.83
Arno, Italy	2,500	0.82
Tiber, Italy	5,000	0.72
Rhone, Lyon, France	5,000	0.33
Rhone, Beaucaire, France	20,000	0.30
Danube, Germany	25,000	0.29
Nile, Egypt (compiled by writer)	....	0.40

To conclude, the foregoing data, limited as they are suggest that variations in rainfall, altitude, or area may each have effects on the "flow variability index" comparable to that ascribed by the authors to "geology." The writer feels that, before the authors' conclusion that "geology" is a principal factor in runoff variation can be accepted, correlation studies must be made to eliminate the effects of variable rainfall, elevation, and area as just pointed out. Although some engineers, lacking other data, may wish to accept the authors' recommendations and estimate unknown stream flow variations following the authors' recommendations of Table 2, the writer believes that, for an unknown stream, there will be less error if the flow variability index be assumed similar to the rainfall variability index, measured or estimated, or if, on the basis of observed physical characteristics of a drainage basis, a "type" variation index be chosen from Table 1.

J. E. McKee,[50] M. ASCE.—It is gratifying to note that Messrs. Lane and Lei have applied standard statistical procedures to stream flow data and have thus analyzed all reliable records of streams east of the 100° meridian. Too often, civil engineers are prone to develop methods and procedures for analyzing data with utter disregard for established statistical parameters which are quite applicable; and as a result they are plagued with hybrid indexes, coefficients, averages, and ratios. The authors have attempted wisely, to avoid such hybrid parameters by using the well-established probability analysis; but (as will be shown) they have apparently overlooked a significant feature of their "variability index."

By its definition and by its method of computation, the variability index must be a logarithm. It is defined by the authors as the standard deviation of the logarithms of the discharge values (X) at 10% intervals, as taken from the duration curve. As such, it is nothing more than the logarithm of the geometric standard deviation, since

$$\log \sigma_g = \sqrt{\frac{\Sigma(\log X - \log M_g)^2}{N - 1}} \qquad (9)$$

[50] Associate Prof. of San. Eng., California Inst. of Technology, Pasadena, Calif.

in which X is a value read from the duration curve; M_g is the geometric mean, since log M_g is the arithmetic mean of the log X values; N is the number of observations (in this case 10); and σ_g is the geometric standard deviation. In the form of its logarithm, the geometric standard deviation has little or no actual significance, other than as another index to be carried in the minds of engineers. As the antilogarithm of the variability index, however, the geometric standard deviation proper has a definite meaning and can be applied directly to stream flow data.

Inasmuch as the duration curves of stream flow data that plot as straight lines on logarithmic probability paper, between limits of 5% and 95%, may be considered as approximately geometrically normal, the geometric standard deviation, σ_g, is the accepted statistical measure of dispersion (or variability) of such data. It has a physical significance in that, when multiplied by, or divided into, the geometric mean (or 50% value in this case), it results in direct values having the same dimensions as the mean. The product and quotient so obtained represent the values at 84.13% and 15.87%, respectively. Thus, if the geometric mean of certain runoff data is expressed as 1.20 cu ft per sec per sq mile and σ_g is found to be 2.20, then the runoff may be expected to be less than 1.20×2.20 or 2.64 cu ft per sec per sq mile for 84.13% of the time and less than $\frac{1.20}{2.20}$ or 0.545 cu ft per sec per sq mile for 15.87% of the time. The variability index used by the authors, which would be log σ_g, or 0.34 in this example, has no such ready application. Few human minds are accustomed to thinking logarithmically, and, hence, few engineers can envision what 0.34 means in this instance.

Although the variability index is definitely established as a logarithm by its definition and by its magnitude in the various tables and examples, and although the characteristic values of the variability coefficient given in Table 2 are logarithms, the authors were apparently thinking of the antilogarithm of their index, or the true geometric standard deviation, when they stated (under the heading, "Use of the Variability Index"):

> "The shape of the duration curve can be obtained by drawing a straight-line duration curve on logarithmic probability paper with a slope such that the ratio of the discharge exceeded 15.87% of the time to the discharge exceeded 50% of the time was equal to the variability index selected."

The quoted statement does not hold for the logarithms used in this paper as the variability index. It holds only for the true geometric standard deviation.

The writer concurs with the authors in their reasoning that a measure of variability of generalized stream flow data is needed and he feels indebted to them for the mass of data that has been thus compiled, analyzed, and compared. He recommends, however, that the variability index be called by its established statistical name—the logarithm of the geometric standard deviation—and that this variability be expressed directly rather than as a logarithm.

G. B. Schroyer,[51] Jun. ASCE, and E. T. Schuleen,[52] M. ASCE.—The authors are to be commended on their excellent attempt to devise a stream flow

[51] Design Engr., Safe Harbor Water Power Corp., Baltimore, Md.

[52] Hydrographic Engr., Pennsylvania Water & Power Co., Conestoga, Pa.

variability index. An accurate appraisal of flow variability, by some kind of rational numerical coefficient, has long been needed in hydrologic investigation, and any step in that direction is salutary. Further commendation is warranted for the authors' astuteness in utilizing, as a basis for their mathe-

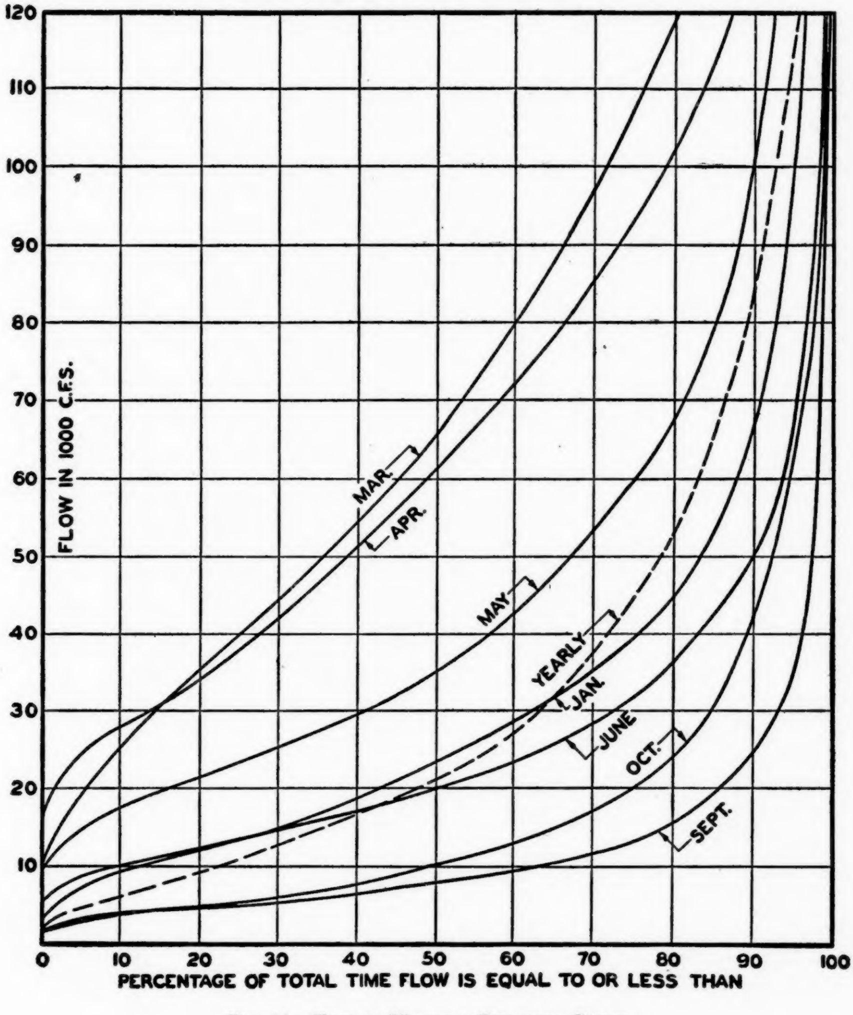

FIG. 14.—TYPICAL MONTHLY DURATION CURVE

matical analysis, the earlier work of the sedimentary geologists, especially W. C. Krumbein, F. J. Pettijohn, and G. H. Otto, as applied in the field of sediment particle analysis. It is often more difficult to visualize the adaptability of work in a foreign field to a given problem than to devise a method of attack on the problem from fundamental concepts of physics and engineering. Similarity of approach to two problems is strikingly apparent only when the two

concepts have been reduced to dimensionless partial differential equations, and a relationship of the factors involved is at an uncomplicated mathematical level. However, as in the case of any new theories, all concepts that are included must be explained.

The three run-of-river hydroelectric plants at Safe Harbor, Pa., Holtwood, Pa., and Conowingo, Md., form an excellent chain of stream gaging stations for the Susquehanna River discharge. Although the three plants have relatively little storage, all flow data have been reduced to natural conditions, and the plants themselves can then be considered as having no effect in equalizing stream discharge and fluctuation in flow. Independent flow records at Holtwood are available from 1917 to date.

It was anticipated that, with such data, valuable confirmation might be offered for the authors' studies of the river; and it is in fact due to this anticipation that the following is presented. For the period from 1917 to 1946, the flow variability index for the Susquehanna River at Holtwood is 0.432. This value differs but slightly from the value of 0.44 determined by Messrs. Lane and Lei for Harrisburg, Pa., and may easily be accounted for by the fact that Holtwood is 47.0 miles downstream from Harrisburg and has a drainage area 2,700 sq miles greater (26,800 sq miles as compared to 24,100 sq miles). Moreover, the geology of the intervening territory is of highly different character from that of most of the drainage basin, since at Harrisburg the river leaves predominantly mountainous terrain and enters rolling farm country. In addition to the yearly duration curve, monthly duration curves for the same period have been prepared and some are depicted in Fig. 14. Using the method outlined by the authors, flow variability indexes for each month were developed from these duration curves. The results, with average monthly flows at Holtwood, are as follows:

Month	Flow variability index	Average flow (cu ft per sec)
January	0.337	33,500
February	0.324	37,400
March	0.306	83,300
April	0.265	74,700
May	0.293	50,500
June	0.270	27,000
July	0.285	18,200
August	0.279	13,600
September	0.324	12,300
October	0.401	18,800
November	0.415	30,900
December	0.376	33,300
Yearly	0.432	36,100

These indexes vary from 0.265 for April to 0.415 for November; they present an unusual condition, although entirely possible in this instance, wherein the annual variability index seemingly bears no relation to the monthly indexes.

Since it can logically be assumed that the geological structure, soil, topography, and the arrangement of tributaries, with regard to time of concentra-

tion of surface flow, are constant for any particular river, the variations in the monthly indexes must arise from one or more other variables. An effort was made to see if such variation could be accounted for by some function of the weather. The monthly variability indexes are plotted in Fig. 15 showing average temperature, precipitation, and runoff. The temperature and precipitation averages are based on records from 33 stations and 58 stations, respectively. It is strikingly evident that no direct correlation between any of these quantities and the flow variability exists; nor is any correlation between the variability index and the runoff coefficient apparent. Temperature appears to be an important factor affecting runoff, and vegetation density would roughly follow the temperature curve.

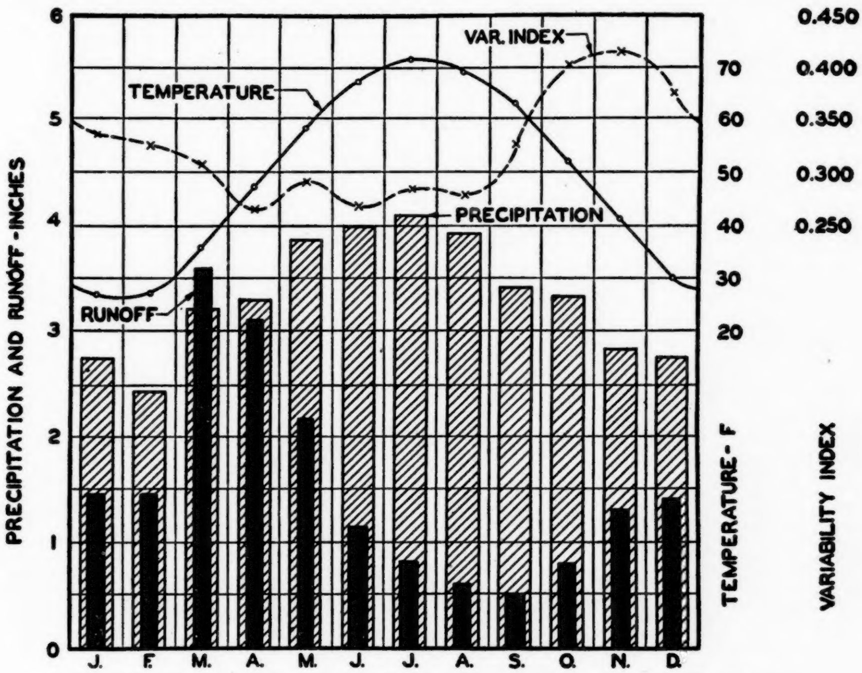

FIG. 15.—MONTHLY PRECIPITATION, RUNOFF, TEMPERATURE, AND FLOW VARIABILITY INDEX FOR THE SUSQUEHANNA RIVER ABOVE HOLTWOOD, PA.

The wide variations in the shape of the duration curves, and, hence, the variability indexes, probably stem largely from the effects of snow and frozen ground in delaying runoff during the winter months and from the higher percentages of flow contributed by ground water during the late summer and fall months. Although the deviations in the monthly variability indexes neither prove nor disprove the authors' contention that geology and topography are the principal factors involved, they act as a warning against the indiscriminate use of indexes to develop duration curves. The foregoing comparison of the relatively constant indexes for April, May, and June, with the widely different duration curves for these months in Fig. 14 illustrates the degree of error possible.

It has been mentioned by Messrs. Lane and Lei that the topography and geology of Pennsylvania are highly diversified, and that extremely varied flow indexes might be expected on individual streams. It should be pointed out that, solely because of this variance, any attempt to predict a variability factor for a given stream would meet with difficulty. As an example of the influence of geology, the authors cite the case of two adjacent small Pennsylvania streams, Muddy Creek with an index of 0.29 and Conewago Creek with a value of 0.58. As the authors mention, Conewago Creek drains an area of Triassic sandstone with traprock intrusions, whereas the drainage basin of Muddy Creek is in a highly laminated and inclined absorbent albite chlorite schist which, were it not for the laminations, would indicate a high index for Muddy Creek. Such unapparent geological considerations as laminations would be completely overlooked in any but the most extensive investigations of a drainage area, and it is very likely that, if such comprehensive geologic data were at hand, adequate stream records also would be available. Moreover, any cursory geological examination would be completely misleading, for, if topography is an indication of the underlying rock structure, identical index values would be predicted for the two streams. The Pennsylvania Gazetteer of Streams[53] lists both streams as being in rolling agricultural country with a narrow main valley flanked by precipitous, well-wooded hills. Another example of topographies not being indicative of the critical geology is mentioned by the authors themselves—that is, the case of Mad River in Ohio. It should be noted also that the period of recorded flow observations for both Conewago Creek and Muddy Creek is of only 10 years duration each; and in addition, in both cases, the periods are during a sustained dry period in Pennsylvania.

Although the variability index does not always reflect certain characteristics of a stream (such as frequency and extent of discharge fluctuations, the equalizing effect of ground water, and other factors peculiar to a given basin), it is hoped that the authors and others will continue in their efforts to place stream flow variation more nearly on an absolute and mathematical basis.

E. W. Lane,[54] M. ASCE.—It is not possible for Mr. Lei to join the writer in preparing the closing discussion of this paper because of the disrupted communication between China and the United States. Although it is believed that he would concur in most, if not all the following remarks, the writer alone is responsible for them.

The large number of constructive criticisms presented by the various discussers has added greatly to the value of the paper and they are appreciated. They fall into two general classes: (1) Those dealing with the statistical aspects of the paper and the mathematics of probability paper, and (2) those dealing with the hydrological aspects and applications of the variability indexes.

Considering first the discussions in class (1), it is advisable to give the reasons for adopting the method of handling the flow of data used in the paper. The data used were drawn from a large number of sources—and were given in four different forms: (1) Plotted duration curves; (2) discharges exceeded during

[53] "Water Resources Inventory Report: Gazetteer of Streams, Pt. III," Water Supply Comm. of Pennsylvania, Harrisburg, Pa., 1917.

[54] Hydr. Consultant on Chf. Engr.'s Staff, Bureau of Reclamation, Denver, Colo.

certain selected percentages of time; (3) percentages of time that the discharge exceeds certain selected values; and (4) mean daily discharges for each day covered by the record. To use all these data, a method which would handle data expressed in any of these forms had to be selected. Since future users of the suggested variability index will also encounter data in all these forms, in order that they may all use the same procedure, it is necessary that the method selected be applicable to all types of data. The method of preparing a duration curve, and of reading values from it at certain selected typical points, was the only process the writer and Mr. Lei could devise which met these requirements.

In making duration-of-flow studies from records of mean daily flow, the shortest method seems to be to divide the entire range of the stream flow into an adequate number of partial ranges and to go through the daily records, tallying the number of days found in each partial range. From the numbers shown by the tallies of these partial ranges the number of days with discharges below the upper limit of each partial range can be easily obtained by adding the tallies in that range to those in all partial ranges covering lower discharges. The writer and Mr. Lei used this method of obtaining the data for plotting the duration curves, where the data used were in the form of mean daily discharges; they believe that any other method is unnecessarily laborious.

Mr. Ownbey expressed the opinion that in computing the index the sum of the squares of the differences of the logarithms from the mean of the logarithms should be divided by 10 instead of by 9, as was done in the paper. Before selecting 9, the writer and Mr. Lei consulted several people who are well versed in statistical methods, in addition to studying the methods of handling similar data recommended by W. C. Krumbein. Still others who are well acquainted with statistical theory have discussed this paper without raising this question; therefore, it seems likely that the use of 9 by the writer and Mr. Lei is sound. If it should not be sound, however, it would have no effect on the practical usefulness of the index developed in the paper, which differs by a constant ratio from the one which Mr. Ownbey would obtain.

Mr. McKee suggests that the geometric standard deviation would be a better value to use for the variability index than the value adopted in the paper. Although the use of the geometric standard deviation as an index would have the advantage of more easy designation, and, therefore, might be more acceptable to statisticians, it would have no practical advantages over the one in the paper, which is the antilog of the one Mr. McKee suggests.

The use of the coefficient of variation, and the use of discharges expressed as ratios to the mean flow have been suggested as possible improvements on the index offered in the paper. One of the advantages of the index used by the writer and Mr. Lei is that, by using the logarithms, the index obtained gives an expression of the relative variability from the mean, and, therefore, accomplishes the same purpose as the coefficient of variation. Converting the discharges to values expressing their magnitudes in terms of the mean and then determining the variability index would result in exactly the same value as was obtained by the method given in the paper. One of the great advantages of the index used by the writer and Mr. Lei is that it is independent of the discharge units used, as Mr. Foster has shown. The values obtained are the same as long as the flows are expressed in units of volume per unit of time.

The writer cannot agree with Mr. Chow's contention that the variability index is a function of the length of the record, when the time is expressed as a percentage of the total time of flow. Of course, there are differences for the duration curves of different years, but this is a chance and not a functional variation.

Several discussers have called attention to an error in the paper (under the heading, "Use of the Variability Index") which resulted from the omission of the word "antilogarithm." The correct statement is:

> "The shape of the duration curve can be obtained by drawing a straight-line duration curve on logarithmic probability paper with a slope such that the ratio of the discharge exceeded 15.87% of the time to the discharge exceeded 50% of the time is equal to the antilogarithm of the variability index selected."

The method given by Mr. Foster for obtaining the predicted flow duration curve from the estimated value of the variability index, and the curve showing the ratio of the median to the mean discharge for various values of the variability index, as given by Mr. Foster and also Messrs. Ospina and Tama, will be of considerable assistance in applying the methods presented in the paper. The explanation of the method of constructing logarithmic probability paper offered by Mr. Foster will be found of great assistance to those who may have occasion to make up such forms.

Turning now to the hydrological aspects and applications of the variability index, Mr. Wing's discussion of the history of the idea of this index is of great interest. As he has stated, its roots go back over seven decades to Francis Galton, and the basic idea was developed by Aime Coutagne about a decade before the independent development of it by the writer and Mr. Lei. Mr. Coutagne expressed his idea in the form of a uniformity index—the higher his value the more uniform the flow—whereas the writer and Mr. Lei developed a variability index, the reciprocal of that developed by Mr. Coutagne, with the higher values allocated to the more variable stream flows.

Some difference of opinion developed in the discussions regarding the variation of the index values with the drainage area. Mr. Lull presents data indicating very little decrease in the variability with drainage area; and Mr. Wing and Messrs. Ospina and Tama believe that the variability decreases appreciably with drainage area. The view that the index decreases with drainage area was given in the paper. However, since the decrease is of a low order (a decrease of perhaps about 50% in index value for an increase of drainage area of 10,000%), it might be argued that the increase was negligible for many cases.

It is regretted that the space available did not permit a more complete justification of the contention that geology and surface storage were the most important factors in controlling flow variability within the area covered by the data submitted. Had the maps showing the variability indexes and their relations to local geology and the discussions of watershed conditions which were a part of the comprehensive paper been published in full, it is believed that this point would have been much more clearly demonstrated. Other factors that have been suggested to account for the variation were rain-

fall variation and forests. In regions of high relief, rainfall variations within short distances can be appreciable, but it is difficult to believe that such can be the case in flat states such as Iowa, Wisconsin, Illinois, and Ohio. In these states variations of 100% or more in the index within short distances were common, and in many watersheds with which this writer was personally acquainted, neither rainfall nor forests seemed to be possible explanations. Mr. Lull gives data which indicate an index about 25% higher for nonforested watersheds than for forested ones in Iowa and Wisconsin. In these states forested areas are allowed to grow timber because they are unsuited to cultivation, either because of high relief, unsuitable soil, or swamps. This is the case in the Iowa streams mentioned, and probably also for those in Wisconsin. These conditions favor low variability for geological reasons, and, therefore, the lower indexes for the forested ones are not necessarily due to the forests.

Perhaps the best arguments for the preponderant influence of surface storage are the very low values of the St. Lawrence, St. Clair, and St. Marys rivers of the Great Lakes system, which are much lower than any others obtained. Strong arguments for geology being very important in the area covered by the records are the low flows of the Kansas and Nebraska streams. Although no quantitative data on the rainfall variability in these Great Plains states are available, there is every reason to believe that it is less uniform than in the states farther east. The watersheds of these streams are nearly treeless. In spite of both these factors, these Great Plains streams give low indexes.

It should have been emphasized more forcefully perhaps that the conclusions of the paper were confined to the areas covered by the flow records available, and that, until data from other regions are available, it is not possible to determine how applicable they are elsewhere. With the exception of some streams in North Carolina and perhaps in southern Illinois, no cases were found which seemed to indicate a rainfall variability effect. However, this writer does not claim that rainfall effects are everywhere generally negligible, but only that they were generally negligible in the area covered by the records. In fact, where there are great differences in rainfall variability there are certain to be differences in flow variability. The flow variability is the result of the interaction of rainfall variability and natural storage, either above ground or below ground; the flow variability varies in the same direction as the rainfall variability and in an inverse direction from the natural storage. It is possible that the low values obtained by Messrs. Ospina and Tama for the streams of the Cauca basin in Western Colombia were caused largely by a comparatively uniform rainfall.

Messrs. Bruin and Hudson have mentioned the necessity of expressing the time in terms of the percentages of time of flow rather than as the percentage of total elapsed time. In all the flows used by the writer and Mr. Lei, these were the same; but in small streams, where the flow ceases a part of the time, it is necessary to use only the time of flow.

Mr. Kazmann has called attention to the lack of exactness in the terms given in Table 2 for permeability, relief, and areas of lakes and swamps. Greater exactness would be desirable, but the labor involved in working up the data on the individual watersheds, in quantitative terms on which generalizations could

be based, would have been prohibitive, even if the data had been available. In most cases this was not true. For classifying the area of lakes and swamps, Fig. 16 will serve as a guide in handling this variable.

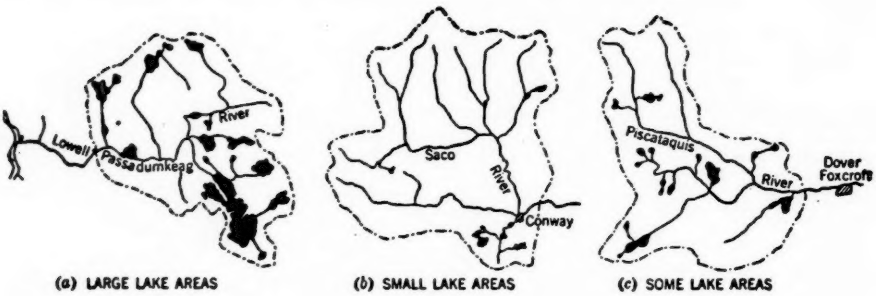

FIG. 16

Messrs. Schroyer and Schuleen have shown the great differences in variability indexes of flow for the Susquehanna River in the different months of the year. This serves to emphasize the necessity of having the data cover one or more complete cycles, such as years, in which the conditions are much the same, rather than to compare conditions which are consistently different, such as those in one month of the year as compared to those in another month.

In judging the geological and surface characteristics influencing flow variability in drainage areas for which geological data are not available, air photographs or airplane reconnaissance would be very valuable. Considerable work has been done toward interpreting these conditions from air photographs,[55,56,57,58] which would repay study by those concerned.

[55] "Airphoto Interpretation of Engineering Sites and Materials," by Jean E. Hittle, *Photogrammetric Engineering*, December, 1949, pp. 589–603.

[56] "Aerial Photographs and Their Applications," by Harold T. U. Smith, D. Appleton-Century Co., Inc., New York, N. Y., 1943.

[57] "The Origin, Distribution, and Airphoto Identification of United States Soils," by D. S. Jenkins, D. J. Belcher, L. E. Gregg, and K. B. Woods, *Technical Development Report No. 52*, Civil Aeronautics Administration, Washington, D. C., 1946.

[58] "Aerial Photographs Used for an Engineering Evaluation of Soil Materials," by R. E. Frost and K. B. Woods, *Proceedings*, 2d International Conference on Soil Mechanics and Foundation Eng., Rotterdam, Vol. I, 1948, pp. 324–330.